Center for Independent Experts (CIE) Independent Peer Review Report on Bering Sea Tanner Crab Stock Assessment Review, 31 July – 3 August 2017

N. G. Hall

Unit 2
2 Wexford Street
Subiaco
Western Australia 6008
Australia
Phone: +61 8 9401 6891
Email: nghall@iinet.net.au

Contents

1. Executive Summary

A Review Panel convened at the NOAA Alaska Fisheries Science Center (AFSC), Seattle, from July 31 – August 3, 2017, to review the stock assessment model TCSAM02, which had been developed for the Bering Sea Tanner crab (*Chionoecetes bairdi*). The Panel found the framework (i.e., mathematical structure) of the current model (i.e., the model that had been demonstrated to be essentially equivalent to the earlier assessment model, TCSAM2013) to be sound, and the software implementation to be very flexible and versatile. Through the use of control files, the specific configuration of the assessment model can be specified at run time, only requiring the writing of additional computer code when the model is extended to accommodate new types of data or processes not already available. Thus, input values of parameters and their priors and bounds, together with parameters to be estimated and likelihood components and weights, can be easily adjusted to refine the model or assess implications of alternative assumptions. The model calculates and outputs OFLs and status determinations, together with an extensive range of tables and links to the R code, which is used to produce the graphical outputs required for model evaluation and comparison.

A jitter analysis using the current base configuration of TCSAM02 (i.e., the model configuration identified as B0 by the Crab Plan Team at its May 2 – 5 meeting at Juneau, AK, and which was being reviewed by the Center for Independent Experts (CIE) Panel) had demonstrated that the model was sensitive to initial values of the parameters, converging (presumably) to local minima of the objective function. Similar convergence issues had also been present in earlier versions of the models used to assess the eastern Bering Sea Tanner crab stock. As in those earlier assessments, the results of the jitter run of TCSAM02 that produced the lowest value of the objective function had therefore been selected as the 'fitted assessment model'. The CIE Panel found that a number of the resulting parameter estimates for this fitted assessment model were at or close to the bounds specified in the control file, suggesting that the model may have become 'trapped' at these locations within the parameter space. The specification of both bounds and priors on individual parameters, together with penalties, weights on associated likelihoods, or the high correlations found among a number of the parameters of the base model, might have been implicated in the convergence difficulties that the model had encountered. It is strongly recommended that, prior to using the model for status determination and informing management decisions, the convergence issues are resolved or, at least, that results of detailed exploration demonstrate that the final result and parameter estimates correspond to values for the global optimum and are not constrained unnecessarily by bounds, informative priors, penalties, or weights on likelihoods. Such constraints may result in bias when determining status or calculating values such as OFLs. Note that a run during the review, using the base model configuration but no bounds, improved the objective function by ~130 likelihood units, but whether parameter estimates remained feasible and how predictions of abundance and size compositions were affected was not assessed.

Convergence issues when fitting length-structured models are not uncommon, as information on stock status and selectivities for fishing fleets and surveys must be obtained from size compositions that reflect mortality and individual growth. This is particularly the case when the difficulties of collecting reliable growth data for Tanner crabs in the eastern Bering Sea are taken into account. The versatility of the TCSAM02 software, and the flexibility that it provides through use of control files, offers considerable ability to explore and identify the basis for the convergence issues, and subsequently resolve those issues, where such ability would have been far less if model configurations had been 'hard-wired' into the software code.

Prior to inclusion of the BSFRF survey data within the TCSAM02 assessment model, it is recommended that an external analysis using a mixed effects model is undertaken of the data from the paired-tow study to determine the relative efficiency and selectivity of the NMFS bottom trawl nets. By capitalising on the side-by-side nature of the BSFRF and NMFS survey tows, such analyses are more likely to elucidate the selectivity curves of the NMFS survey nets than would result by simply including the BSFRF data in TCSAM02 as an additional survey, with only partial coverage of the spatial extent of the area occupied by the eastern Bering Sea Tanner crab population. Such external analysis will assist in determining how best, subsequently, to include the BSFRF survey data within TCSAM02.

Collection of growth data from the eastern Bering Sea population of Tanner crabs should be continued with the aim of increasing sample sizes and improving the growth models while ensuring, to the extent possible, that the Tanner crabs employed in the growth study are representative of the individuals in the population.

TCSAM02 is the best stock assessment model currently available for determining the status of the eastern Bering Sea Tanner crab population. It makes better use of the data that are available than its predecessor, and, by using growth data from the eastern Bering Sea population rather than from the Kodiak region, provides an assessment that is likely to be more reliable. As with earlier assessment models for this crab stock, the TCSAM02 model fails to converge to a global minimum, and relies on jittering of input parameters to conduct a random search of the likelihood surface and select a minimum from the results of that search for use in status determination and OFL calculation. It is suggested that, for results to be considered reliable, it is important that the convergence problem is overcome such that TCSAM02 converges consistently to the global minimum.

2. Background

2.1. Overview

A Review Panel convened at the NOAA Alaska Fisheries Science Center (AFSC), Seattle, from July 31 – August 3, 2017, to review the stock assessment model TCSAM02, which had been developed for the Bering Sea Tanner crab (*Chionoecetes bairdi*). The agenda for the Review Workshop is presented in Annex 3 of Appendix 2.

The Statement of Work provided by the CIE is attached as Appendix 2. This CIE report is prepared in accordance with the requirements of this Statement of Work, and describes the assessment and review process.

Prior to the review meeting, the documents describing the assessment model and other background material had been made available to Panel members. A list of these and other documents received during the review is presented in Appendix 1.

2.2. Terms of Reference

The terms of reference for the review of the stock assessment model TCSAM02 are presented in the Statement of Work (Appendix 2).

2.3. Panel membership

Details of the Panel Membership and of other key participants for the review of the stock assessment model TCSAM02 are presented in Appendix 3. In particular, the review was chaired by Dr Martin Dorn and the CIE Reviewers were Drs Cathy Dichmont, Anders Neilsen, and Norman Hall.

2.4. Date and place

The Review Panel convened at the NOAA Alaska Fisheries Science Center (AFSC), Seattle, from July 31 – August 3, 2017, to review the stock assessment model TCSAM02, which had been developed for the Bering Sea Tanner crab.

2.5. Acknowledgments

Thanks are expressed to the various individuals who participated in the review meeting, and who contributed their knowledge and advice, for making the review such an interesting and positive experience. Dr Stockhausen is to be commended for the quality of the assessment model framework that he has developed, for the extensive background material he provided, and for his very competent and professional responses to the Panel's queries and requests. Thanks are also extended to Dr Dorn for his excellent chairmanship of the meeting.

3. Description of Reviewer's role in review activities

Prior to the review meeting, I downloaded and familiarised myself with the background documentation, the description of the model that was the subject of the review, and the associated model inputs and outputs for various configurations of this model (Appendix 1). Subsequently, I attended and actively participated as a Panel member in the CIE Review that was held in Seattle, and prepared this report.

4. Summary of findings relevant to the Terms of Reference

During the review meeting, the various participants presented details of the biology of the Tanner crab (*Chionoecetes bairdi*) population in the eastern Bering Sea, the surveys that were used to collect fishery independent abundance and composition data, the fishery and its management, the stock assessment and projection model, and the research that was currently being undertaken. Throughout these presentations, the Review Panel requested clarification or sought additional detail from the presenters. Given the changes in the names used to designate the same models in different reports, the participants were asked to identify the particular model that was the 'current assessment model', which the Statement of Work required the CIE reviewers to review. The Panel also requested a table summarizing the time blocking that had been employed in this model, together with a number of runs intended to facilitate determination of the basis for the convergence issues that had been identified.

ToR 1. Statements assessing the strengths and weaknesses of the current Tanner crab stock assessment model with regard to population dynamics, fishery and survey components, likelihood components, and model evaluation.

Overview of TCSAM02

The current assessment model is identified by the Crab Plan Team in the report from the May 2017 CPT meeting as 'B0: the matching model using TCSAM2', i.e., the model identified as T02A in Stockhausen's (2017) Tanner Crab Assessment Report for the May 2017 CPT Meeting. The code for this model, i.e., TCSAM02, was provided to the CIE review meeting in the directory, 'basemodel'. This model describes the dynamics of the Tanner crab population in the eastern Bering Sea from 1948 to 2016, with model years that run from July 1 to June 30. The model 'recruits' small, immature, new-shell crabs to the population at June 30/July 1, while (directed and bycatch) fishing and molting/growth/maturation are represented as discrete events that occur at specified occasions within the annual cycle. The directed Tanner crab fishery retains only male crabs of a preferred size, releasing all females.

At the time of the CIE Review, the model was found to be encountering convergence issues. Jitter analysis had demonstrated that the model converged to different locations in the parameter space when different initial values of the parameters were applied, and that some of these jittered results considerably improved on the fit that had first been obtained. As an interim solution, the best of the jittered model fits had been employed as the best fitting model when producing model output, status determination, and management advice. Unfortunately, there is no guarantee that the current best fitting model corresponds to the global optimum. Indeed, examination of parameter estimates indicated that ~10 of the parameters were at or very close to their bounds, suggesting that the model might possibly have been 'trapped' at these locations in parameter space.

Strengths:

The single-area, single-species, size-based model, TCSAM02, which describes the dynamics of the Tanner crab (*Chionoecetes bairdi*) in the eastern Bering Sea and the fisheries associated with this crab population, extends earlier stock assessment models for this species. The new model accommodates a variable number of fisheries, surveys, and likelihood components, time blocks for parameters, and alternative selectivity functions. The flexibility of the model framework is achieved through use of a control file that specifies the configuration of the model to be used when the model is run and avoids the need to modify the computer code and recompile the model. Such flexibility within the underlying model framework readily allows exploration of different model configurations and different options, such as the data to be employed in the analysis, the parameters to be fixed or estimated, the bounds or priors on those parameters that are estimated, penalties, and weights to be applied to component likelihoods. By avoiding the need to modify and extend the computer code within the model, introduction of coding errors is avoided and, through use, the model framework becomes increasingly well-tested thereby expediting the process of model development.

TCSAM02 is viewed as a stepping-stone towards future inclusion of Tanner crab in the GMACS model, initial development of which was focused on the red king crab. Biological and fishery differences, e.g., the implications of the terminal molt undergone by both sexes of Tanner crab, currently preclude such inclusion. The development of TCSAM02 resolves many of the issues that will ultimately need to be addressed when extending GMACS to include assessment of Tanner crabs.

As demonstrated by the scatter of stock-recruitment points estimated using the 2012 assessment model (in the presentation of the biology of the eastern Bering Sea Tanner crabs), it would be difficult to produce a reliable description of a stock-recruitment (S/R) relationship for this stock. Accordingly, the decision to employ a Tier 3 assessment appears appropriate. Such an assessment is employed when it is not possible to determine a reliable S/R relationship, but proxies for F_{MSY} and B_{MSY} (i.e., $F_{35\%}$, $B_{35\%}$) can be estimated.

Weaknesses:

The failure of the model to converge to a global minimum for the overall negative log-likelihood is of concern. A number of analyses were undertaken during the course of the review to investigate this convergence issue. While these provided approaches that might ultimately assist in resolving the problem, a solution was not found during the meeting. The Panel suspected that some of the convergence issues might reflect the mix of bounds, priors, and penalties on parameters and the weights on likelihood components and advised that, as a first step towards resolving the problem, these constraints should be removed. One approach that could be employed was the use of transformed parameters. For example, use of a parameter to be estimated such as the 'natural logarithm of natural mortality' ensures that natural mortality, i.e., exp(log of natural mortality), is positive. A plot of the negative log-likelihoods obtained when jittering TCSAM02 suggested that the likelihood surface might reflect a number of states, such as might result from the high correlation found to exist among a number of parameter estimates, confounding of parameter estimates, or tension among data sets.

Likelihood profiles for TCSAM02 examined during the review possessed extremely high and very unusual peaks. Subsequent to the review, however, similar anomalous peaks in likelihood profiles were found for ADMB models other than the Tanner crab assessment model. These appear to have resulted from a 'glitch' in the most recent version of ADMB and/or its compiler, and the issue has now been referred to and resolved by the ADMB developer group.

With TCSAM02 representing the eastern Bering Sea Tanner crab population as a single stock, temporally-varying changes in size composition and spatial distributions of the crabs of different sizes and maturity states are likely to introduce changes in selectivity for the directed and bycatch fisheries. The model must attempt to match the observed data and selectivity is a key variable that informs these predictions. Currently only two time blocks are used for selectivity of both the survey and fishery data. It may be useful to consider increasing the number of selectivity time blocks that are used, as the additional number of parameters involved for each time block is relatively small. The possibility of relating selectivity to temperature might be investigated, as this appears to be an important factor affecting growth and spatial distribution of immature and mature female and male crabs.

The separation of Tanner crabs of the eastern and western Bering Sea by a line at 166° W is arbitrary. Biological characteristics, e.g. abundance, size distribution, and distributions of male and female juveniles and adults, vary within regions both spatially and temporally, and suggest different patterns of movement of female and male Tanner crabs. Improved knowledge of factors affecting movement, growth, and distribution of immature and mature females and males would allow the implications of these factors for TCSAM02 to be assessed.

The measure that is used as an index of reproductive potential for the Tanner crab population is mature male biomass, MMB, not egg production or a proxy for this such as mature female biomass. Sexually-dimorphic growth and a fishery that exploits only male Tanner crabs, coupled with low selectivity of female Tanner crabs, make it difficult to accurately assess the mature female biomass and use this as an index of reproduction. The nature of the relationship between MMB and reproductive potential is further complicated by the ability of multiparous females to use stored sperm to fertilize egg clutches. The appropriateness of MMB as a proxy for reproductive potential requires further research. Data on the proportion of females that are barren have been collected in the past by NMFS. The development of a time series of such data, and inclusion in the model of predictions of such proportions, might provide a valuable link between MMB and female reproductive potential. Although the comment was made at the review meeting that there was no indication that a lack of males was impacting the ability of the females to reproduce, data to support this assertion were not presented.

It is assumed that, following recruitment to the modelled stock at June 30/July 1, all crabs molt once a year at a specified time until they undertake a terminal molt. It is recognised, however, that younger juveniles may undertake multiple molts in the first two to three years of life. No evidence is presented to demonstrate the validity of the assumption of a single molt per year for all new shell crabs.

The conclusion drawn by Tamone et al. (2007) that, because of their low ecdysteroid levels, large-claw Tanner crabs, i.e., those with CH/CW>0.18, have negligible potential for molting is inconsistent with the evidence of Paul and Paul (1995) that many such males do in fact molt, with molt intervals exceeding 1 year. The assumption in the model that male Tanner crabs have a terminal molt should be revisited.

The model assumes an increased natural mortality of mature crabs between 1980 and 1984 in order to improve the quality of fit. Justification for this assumption and explanation is poor, e.g. 1980–1984 "has been identified as a period of enhanced natural mortality on mature crab, the mechanisms for which are not understood". What data provide evidence of enhanced natural mortality? Are other explanations possible?

Hybridization is currently ignored by the model, yet may represent a loss of female reproductive potential of 'pure' Tanner crabs if their females are mated by male snow crabs (*Chionoecetes opilio*) or snow-Tanner crab hybrids. Note that it was advised during the review meeting that Tanner crab fishers are required to discard hybrid crabs and species other than Tanner crab.

While TCSAM02 has been developed to determine the current overfishing limit (OFL), acceptable biological catch (ABC), and annual catch limits (ACLs), and whether the stock is overfished and/or overfishing is occurring, it is not clear that it currently allows exploration for the State of Alaska (SOA) of the effects of different size limits, gear restrictions, and fishing seasons, i.e. information that would assist the State's management process.

Model nomenclature is inconsistent, changing from one report to the next.

ToR 2. Statements assessing the strengths and weaknesses of the current Tanner crab stock projection model, with regard to methodology.

Approach used in TCSAM02

Projections assume that selectivity and retention in the directed fishery are those estimated for the most recent assessment year, whereas selectivity curves for the bycatch fisheries are assumed to be the averages for the last five years. Fishing mortalities of Tanner crabs for the Bristol Bay red king crab and groundfish fisheries are set to the average fishing mortalities for the last five years. The fishing mortality on Tanner crabs for the snow crab fishery is set to the average fishing mortality on Tanner crabs for the last five years scaled by the ratio of the F_{OFL} for snow crabs to the average F for snow crabs over the last five years. Recruitment is set to the average from 1982, with initial population numbers at July 1 set to the values estimated at the end of the last year by the assessment model. F in the directed fishery is calculated as the fishing mortality that results in a long-term equilibrium mature male biomass-at-mating, which is 35% of that associated with zero fishing mortality ($B_{35\%} = 0.35B_0$). This fishing mortality is the estimated $F_{35\%}$ for Tanner crabs. F in the directed fishery is then set to $F_{35\%}$ and current MMB at Feb 15 is calculated by projecting from the initial numbers at the end of the assessment period. If MMB exceeds $B_{35\%}$, $F_{\mathrm{OFL}} = F_{35\%}$, otherwise a new value of F is calculated using the harvest rule and the ratio of current MMB to $B_{35\%}$ and a new estimate of current MMB is calculated. This latter iterative procedure is repeated until MMB reaches equilibrium, at which stage F_{OFL} is set to the final estimate of F. The OFL is then calculated as the total catch biomass (directed plus bycatch) estimated by projecting one year beyond the last assessment with F in the directed fishery set to F_{OFL}.

Strengths:

The inclusion of OFL calculations within the assessment model improves on the previous use of a projection model separate from the TCSAM2013 assessment model, as it allows propagation of the uncertainty of the assessment into the estimate at equilibrium of $B_{35\%}$ and of the uncertainty of the OFL when current MMB for $F = F_{35\%}$ exceeds $B_{35\%}$ or when MCMC chains are generated (noting that automatically-differentiated (AD) variables are not used in iterative calculations of OFL).

ToR 3. A review of the fishery dependent and independent data inputs to the stock assessment with regard to quality of information and appropriateness to the assessment.

Stock assessments of the eastern Bering Sea Tanner crab population are well informed by the biological and research data that are available. In particular, the long time series of very informative and detailed abundance and compositional data from the Bering Sea Bottom Trawl survey, conducted annually by the NMFS, provides essential data that facilitate assessment. As with other crab populations, however, the absence of a reliable method for determining the ages of individual crabs requires the development of assessment models that, through use of growth data, describe the dynamics of the stock in terms of the changes in length compositions over time.

Strengths:

The trawl surveys conducted in the eastern Bering Sea by NMFS from 1975 have produced valuable time series of fishery-independent abundance and size composition data that provide the model with the key information on the trends in relative abundances of the different categories of Tanner crabs and mortality that it requires.

The growth data from the eastern Bering Sea, which are now used in the model, are more representative of the Tanner crab population to which TCSAM02 is applied than the growth data collected from the Kodiak region. Even so, the selection of crabs for inclusion in growth studies appears likely to be non-random. Given that the relationships between proportions of mature males and carapace width appear (in Powerpoint presentations at the review meeting) to vary both spatially and temporally, there is potential that growth also varies or that, through movement, the rate of growth or temperature may influence the spatial distribution of Tanner crabs. It is therefore appropriate that growth parameters are estimated within TCSAM02 rather than externally, as this ensures that the fitted relationships between molt increment and pre-molt carapace width are also consistent with size composition and abundance data and with the population dynamics described by the model.

Deviations of the ‘observed’ proportions of new shell male Tanner crabs molting to maturity at different carapace widths from the proportions predicted using the fitted mean maturity ogives were relatively minor.

Despite differences in biological characteristics and abundances of the different categories of crabs in the areas east and west of 166° W, the decision to model the population using a single-area model is justified due to the paucity of data on movement, and contributions of egg production from each region to subsequent recruitment to the different regions.

Although the annual spatial distributions of water temperature are monitored, and are presumably a key factor in determining the spatial distribution of Tanner crabs and a key influence on biological processes such as growth, temperature is not directly taken into account in the TCSAM02 framework. At this stage of model development, however, it would be inappropriate to increase the model's complexity by inclusion of this variable, particularly as its influence on biological processes is not yet sufficiently well determined.

Input sample sizes for NMFS survey data are set to 200, while those for fishery size compositions are set to the ratio of number of crabs measured relative to the average number measured per vessel, up to a maximum of 200. Such input sample sizes appear appropriate, recognizing that these are typically adjusted subsequently using an approach such as that proposed by McAllister and Ianelli (1997) or Francis (2011).

Weaknesses:

Catch data for the foreign fishery fleets are imprecise, and discard and size composition data are unavailable. Due to the lack of size composition data for the earlier years of the historical time series of Tanner crab catches, a common selectivity curve based on US data has been employed. Discard data were also unavailable for the earlier years of the fishery. The retention curve for the first time block considered by the model was therefore that determined for the US groundfish fishery. Sensitivity of the model to alternative time series of data should be explored to assess the implications of the uncertainty of the early data. A run using higher foreign catches was undertaken during the course of the review meeting to ascertain whether such catches might have affected the estimates of the higher levels of natural mortality experienced in the early 1980s. Although the mortality estimates for this apparent 'high mortality' period were slightly reduced, underestimation of foreign catches was unlikely to offer an alternative explanation for the data from 1980—1984.

Growth data for Tanner crab in the eastern Bering Sea remain limited, with molt increments available for only 68 males and 57 females. Collection of additional growth data to increase sample size, extend the study over additional years, and obtain a greater spatial coverage of the population should continue such that relationships between molt increment and pre-molt width can be refined further, and data for each sex become available for a broader range of pre-molt carapace widths, and, in the case of males, a broader range of pre-molt chela heights. Such data will also improve estimates of the relationships between the probabilities of undertaking a terminal molt and pre-molt carapace width for female Tanner crabs from the eastern Bering Sea and molt to morphometric maturity for male Tanner crabs.

From the descriptions provided in the background material, it was not possible to determine whether selection of Tanner crabs from the eastern Bering Sea population for inclusion in growth studies was opportunistic (and thus non-representative and potentially biased) or whether it followed a well-designed sampling protocol and was likely to produce representative growth data. When sampling, attention should be paid to factors likely to affect growth, such as temperature and location of sampling, and data relating to these factors recorded for use in future analyses.

In the case of the male Tanner crabs, data on chela height (CH) and carapace width (CW) collected from Glacier Bay were used by Tamone et al. (2007) to determine the frequency distribution of CH:CW ratios. This distribution was found to be bimodal, with a minimum frequency of crabs between the two modes at a ratio of 0.18. This value has subsequently been used to classify male Tanner crabs as immature if the ratio CH:CW < 0.18 and mature if CH:CW > 0.18. Subsequent analyses of NMFS EBS bottom trawl survey chela height and carapace width data for males described by Stockhausen (2017; "NMFS Survey Data: Chela Height Data and Male Maturity Ogives") demonstrates that use of this criterion classifies many "old shell" males as immature. Such a result is inconsistent with the view that immature males molt each year and should therefore be observed as "new shell" crabs, although the possibility that the immature "old shell" males have skipped a molt is also feasible. Accordingly, in subsequent analysis it has been assumed that the ratio 0.18 provides a valid criterion for classification of maturity status for male Tanner crabs from the eastern Bering Sea and immature "old shell" males have henceforth been treated as immature "new shell" crabs. It is possible, however, that the CH:CW ratio of 0.18 is peculiar to the sample of male Tanner crabs collected by Tamone et al. (2007) from Glacier Bay, and that samples collected from different regions of the eastern Bering Sea and different years might differ in CH-CW allometry. It is recommended that, to test the validity of the CH:CW criterion of 0.18, the natural logarithms of chela height and natural logarithms of carapace width for males from different regions of the eastern Bering Sea and different years be subjected to linear discriminant analysis, and the resulting maturity classifications compared.

The estimate of natural mortality, M= 0.23 year^{-1}, the natural logarithm of which is input to TCSAM02 as the 'base_ln-scale_M', is based on an estimate of maximum age and, as Tanner crabs cannot be aged, is very uncertain. This should not be an issue, however, provided the available data are informative and the specified prior distribution for natural mortality is uninformative, as multipliers of this natural mortality are estimated for immature crabs, mature males and mature females. Multipliers are also estimated for increased natural mortality between 1980 and 1984.

Although input sample sizes are typically adjusted using an iterative reweighting approach such as that proposed by McAllister and Ianelli (1997) or Francis (2011), at the current stage of development of TCSAM02, no reweighting has been undertaken.

The decisions regarding discards made by the fishers are likely to be influenced by market demands. Observers determine subjectively whether a fisher would retain or discard an individual crab, but their decision may differ from that which a fisher might make when sorting catches. This is reflected in a mismatch between the size composition of retained crabs measured dock side and those that are recorded as 'probably retained' by the observer. More accurate retention data would assist the model in determining the retention curve and providing more accurate estimates of discard catches.

NMFS EBS bottom trawl survey methods initially employed in 1975 were progressively standardised and measurements of area swept improved over time. Prior to 1980, no net mensuration gear was used in the NMFS survey. In 1982, the net used was standardised, with the 83-112 trawl being used by all vessels between 1982 and 2013. Improved net mensuration gear was introduced for the NMFS survey in 1987. From the equations presented at the review meeting, it appears that the estimates of abundance (and associated imprecision) calculated from the counts of crabs within the different grid cells and areas swept in those cells assume that values of areas swept are equally precise throughout the entire time series. To assess the implications of the relatively poorer quality of the estimates of area swept in the earlier years of the NMFS survey data, a sensitivity trial could be run with the CVs for the estimates of abundance for those earlier years increased by an appropriate amount.

The spatial extent of the NMFS EBS bottom trawl survey changed over time, with apparently incomplete coverage of the spatial extent of the distribution of Tanner crab in the eastern Bering Sea prior to 1979. The time series of NMFS survey data that is input to the base model configuration includes estimates from 1975 through 1978. It is not apparent from the input CVs that the data for these earlier years were adjusted to account for incomplete coverage.

ToR 4. Recommendations for alternative approaches to evaluate model convergence and compare multiple models.

Convergence

Convergence in ADMB when minimizing the objective function is typically assessed using the criteria that (1) the Hessian is positive-definite and can be inverted to estimate the variance-covariance matrix for the parameter estimates; (2) the maximum gradient is less than the critical value that has been specified by the 'convergence_criteria' statement in the 'RUNTIME_SECTION' of ADMB (or the default value for this variable); and (3) jitter analysis demonstrates that the model consistently returns to the same point in parameter space with the same value of the objective function.

The issue of model convergence was a key consideration of the Review Panel during the meeting. Jitter analysis is typically used to assess whether or not a model successfully converges to the same point in parameter space and on the likelihood surface. If such convergence is not demonstrated, the structure of the model is usually re-examined and tensions among different data components resolved, or phases of parameter estimation revised, such that the model ultimately converges satisfactorily. In the case of the TCSAM02 model, the jitter analysis has been treated as a further element of the model fitting process, with the final 'best fitting' model selected as the jitter run that minimized the objective function, but without assurance that the value of the objective function lies at the global minimum. Various recommendations were made by the Panel to aid in locating the basis for the failure of TCSAM02 to converge to a global minimum. These included some of the following suggestions.

Fishery assessment models are typically nonlinear and complex, and thus are often difficult to fit. Convergence problems are not unusual, and reflect the interaction of model structure, data, and minimization algorithm. Convergence should not be assumed, but should be demonstrated by examining the results of diagnostic outputs (e.g. likelihood profiles) and exploratory analyses (e.g. jitter and residual analyses). The nature of the likelihood surface, which is a function of model structure and data, will determine whether a particular minimization algorithm is able to converge from some initial point within the region of the parameter space in which parameter values are feasible to a minimum within that range and, if so, whether that minimum is a local or global minimum in that region. By default, ADMB uses a quasi-Newton approach to minimize the objective function defined by the model. If this fails to converge to a minimum, an alternative derivative-free algorithm, i.e., the Nelder-Mead simplex algorithm, may be applied by using an ADMB command line option (see ADMB user manual).

The selection of initial values of parameters is often important when attempting to fit nonlinear models. The iterative minimization algorithms typically require that this initial point in parameter space must lie within the neighbourhood of the minimum, such that the gradients of the objective function with respect to the estimated parameters inform the algorithm of the direction in which parameter values should be moved in order to locate the minimum.

If ADMB converges, the Hessian should be positive definite and the maximum gradient should be small. If not, it is possible that the objective function is not differentiable with respect to all the parameters, a parameter has hit a bound (or is essentially bounded by the form of its prior), one or more parameters are highly correlated or confounded, or the value of the objective function is independent of the value of a parameter. Removal of sdreport variables may sometimes assist if the Hessian is not positive-definite, and these variables are poorly defined. In some cases, correlation among parameters may be addressed by re-parameterisation, e.g. use of lengths at two reference ages rather than asymptotic length and Brody-Bertalanffy growth coefficient in the von Bertalanffy growth curve. Another form of re-parameterisation, i.e. re-scaling, may be used to aid model convergence when parameters vary markedly in scale. The re-scaling may sometimes be achieved by transforming observed data to different units of measurement.

In the case of TCSAM02, ADMB converged to a location in the parameter space at which approximately 10 parameters were at or very close to the bounds that had been specified. These bounds may have been intended to constrain the fitted parameters to a feasible range of parameters, such that a parameter value at the boundary is essentially very unlikely. It is possible, however, that during its search the minimization algorithm moved the current point in parameter space to a location from which escape was impossible. By appropriate phasing of the different parameters to be fitted, it may be possible to avoid such capture. Further information may be obtained by fixing selected parameters to their initial values (by setting the phase at which they are to be estimated to a negative value), and thereby reducing model complexity. Bounds that are unnecessary are typically removed in the final estimation phases as these may influence estimates of uncertainty. Bounds may also be removed by specifying parameters in terms of transformed values. For example, use of the log of a parameter p as the parameter to be estimated, which ensures that p=exp(log of p) is positive, avoids the need to specify the lower bound of p to a value slightly greater than zero when using p as a parameter.

If a prior for a parameter is specified, the posterior distribution for the parameter should be compared with its prior distribution, to ensure that the prior is not informative. If the prior is informative, it is likely to constrain the value of the estimate that is obtained when ADMB is run. Similarly, penalties and weights on likelihood components will constrain the values of the parameter estimates.

Other approaches that are often used in resolving or understanding convergence issues include the production of likelihood profiles (for the overall likelihood and the various likelihood components) for key parameters and derived variables. Such profiles provide information regarding the tensions among the different datasets and the values of other parameters that are favoured by each dataset. When calculating the profiles, values of fitted key parameters may be stored together with the values of the overall likelihood and the likelihood for each of the component datasets. The resulting data may be plotted to assist in determining the nature of the likelihood surface, and thus providing information that may assist in resolving convergence issues, e.g. whether there are local minima at which the parameters become ‘stuck’. Tension among datasets may relate to either model structure or model predictions that fail to represent the observed data, possibly requiring revision of assumptions and model code.

Another approach to resolving convergence issues is to increase model complexity incrementally by introducing datasets successively in an order that matches the extent to which the data are considered to be most representative of the modelled population. The resulting changes in parameter estimates and in the contribution of the different likelihood components that result from such incremental increase in complexity provide information on the nature of the likelihood surface and how the different datasets affect the location of the parameter estimates. An alternative to incremental increase in complexity is an approach that reduces complexity. Starting with the current, full model, each dataset is removed and the model refitted and the effect of that removal on likelihood and parameter estimates is examined and compared. Additional datasets may then be removed and the process repeated. Through such exploration, it is often possible to identify trade-offs among different assumptions and to identify the issues within the model assumptions, code, or data that have produced the convergence issues.

Independent implementation of the model in another computer language or software package or simulation of data using the population dynamics employed in the model may assist in identifying problems that are leading to convergence issues. MCMC or Sampling-Importance-Resampling analyses (the latter undertaken outside ADMB) may assist in identifying convergence problems.

It should be recognised that there may be cases in which the information contained within the data are insufficient to provide a clearly defined minimum. In such cases, the likelihood surface is likely to be fairly flat and jitter tests will produce estimates where parameters are located at different points in parameter space, but the value of the objective function remains relatively constant.

Model convergence is typically tested using jitter analysis, by ensuring that the Hessian is positive definite, by a gradient of the objective function relative to each parameter that is sufficiently small, by examination of likelihood profiles, and by examination of residuals for each dataset and of retrospective patterns.

Model comparison

Model comparison typically involves examination of the log-likelihoods for the different datasets and the values of key parameters to determine the trade-offs involved and to identify where model fit has improved or become degraded. If the total number of parameters used in the various models differ, changes in overall likelihood are likely to be as expected. It is my understanding that participants at the data weighting workshop by CAPAM considered use of AIC to determine the best-fitting model inappropriate. It should be recognised that the results produced by the different candidate models reflect model uncertainty, and thus sensitivity of model results and predictions to alternative assumptions regarding population dynamics and data. Although different models will result in different status determinations and OFLs, selection of the base model should be based on the extents to which the assumptions of the different models are justified and the quality of the fits to the different datasets used by the models, not subjective decisions based on the predictions of spawning biomass, recruitment deviations, and values of variables of interest to fishery managers.

ToR 5. Recommendations for integrating BSFRF surveys into the assessment.

Use of BSFRF surveys in other crab assessments

The snow crab assessment treats the BSFRF data as an additional independent survey, and assumes that, for all size classes in the assessment model, the catchability of the BSFRF nephrops gear is 1. Thus, if $n_{y,z}$ is the population abundance in year y and size bin z, in the area jointly sampled by the BSFRF and NMFS trawl surveys, and $n_{y,z}^{\mathrm{BSFRF}}$ is the estimated number of crabs of that size in the jointly sampled area based on the catches of the BSFRF trawl gear (presumably expanded from area swept to the area that was jointly sampled), then $n_{y,z}^{\mathrm{BSFRF}} = A_{y,z} n_{y,z}$, where $A_{y,z}$ is the annual, size-specific availability of crabs of that size in the jointly sampled area. Then, if $n_{y,z}^{\mathrm{NMFS}}$ is the estimated number of crabs in that same jointly sampled area produced by expanding the catches of the NMFS survey from area swept to jointly-surveyed area, $n_{y,z}^{\mathrm{NMFS}} = q^{\mathrm{NMFS}} S_z^{\mathrm{NMFS}} n_{y,z}^{\mathrm{BSFRF}} = q^{\mathrm{NMFS}} S_z^{\mathrm{NMFS}} A_{y,z} n_{y,z}$, where q^{NMFS} and S_z^{NMFS} represent the relative catchability of the NMFS trawl shot relative to the BSFRF trawl shot and selectivity of crabs of size bin z, respectively. Both $n_{y,z}^{\mathrm{BSFRF}}$ and $n_{y,z}^{\mathrm{NMFS}}$ are considered to have a log-normal distribution about the above expressions. A similar approach has not been employed in the Tanner crab assessment model due to concern that the BSFRF side by side surveys in 2013–2015 only covered Bristol Bay, not the entire spatial extent of the eastern Bering Sea Tanner crab population.

Strengths:

By fishing side-by-side with the vessels undertaking the NMFS trawl survey in the Bristol Bay region. the BSFRF trawl survey, which used a small-mesh Nephrops net, has undertaken a paired-tow study that has produced valuable data relating to survey selectivity and efficiency. The information on the selectivity of the NMFS trawl obtained from appropriate mixed effects analyses of these data should assist in refining the Tanner crab assessment. It is recommended that such analyses (e.g. Fowler & Showell, 2009; Cadigan & Dowden, 2010, Kotwicki et al., 2017) be undertaken independently of the assessment model in the first instance, incorporating the approach within the model at a later stage when all issues of the analysis have been resolved.

Weaknesses:

Use of the BSFRF data as an additional fishery-independent survey, as in the snow crab assessment, would fail to make use of the side-by-side nature of the NMFS and BSFRF surveys. For the snow crab assessment, Dr Noel Cadigan had recommended the use of a mixed effects model to compare the catches made in the paired tows, thereby estimating the selectivity of the 83-112 Eastern trawl survey net relative to that of the Nephrops trawl (e.g. Fowler & Showell, 2009), and had suggested that survey grid cell might be employed as a factor. For that assessment, Dr Cadigan had also advised that the assumption that the catchability of the Nephrops trawl is equal to 1 might be invalid.

ToR 6. Recommendations for alternative assessment/projection model configurations.

Consider constructing a simple biomass dynamics model of the fishery using the catches and bycatches that are removed from the Tanner crab stock and the indices of abundance provided by the NMFS bottom trawl survey. Such a model would provide a broad indication of the population dynamics of the crab population and identify the main trends in relative abundance. It provides a useful basis for comparison with the results that are subsequently obtained by using the additional size composition data, which will presumably add information regarding scale.

Temperature appears to play a major role in determining the growth and spatial distributions of the immature and mature female and male Tanner crabs. It would be useful to consider how temperature might influence factors such as growth and selectivity, and then, by appropriate extension of the model, explore whether introduction of temperature results in improvement of model fit.

At a later stage, consideration might be given to moving the model from ADMB to Template Model Builder (TMB), and subsequently describing the fishery using a state-space framework.

ToR 7. Recommendations for research that would reduce the uncertainty associated with key parameters assumed or estimated in the assessment.

A key assumption of the assessment model for the eastern Bering Sea Tanner crab population is that Mature Male Biomass (MMB) at mating is an appropriate proxy for the reproductive potential of the stock, and that reference points for this variable are the same as those that are typically used for mature female biomass at spawning. This assumption should be the subject of research, possibly using a management strategy framework to explore the robustness of harvest control rules and management approaches based on MMB.

Although ADMB will produce estimates of parameters that produce the best-fitting assessment model given the specified likelihoods, bounds, priors, penalties and weights, simulation studies should be undertaken to demonstrate that the model is converging to the correct location in parameter space and that it is not producing biased parameter estimates. The resulting estimates should be of sufficient precision to ensure that management advice is robust.

The data that are currently available on the growth of the female and male Tanner crabs from the eastern Bering Sea are limited, and it is unclear that the growth data that have been obtained are representative of the growth of the entire population and that molt frequencies are as assumed. Further research on growth (and molt to maturity) is required, but consideration should be given to ensuring that, to the extent possible, representative data are collected. Factors affecting growth should be investigated.

As a longer term initiative, the movement and distribution of immature and mature female and male Tanner crabs, and factors affecting that movement and those distributions, should be investigated.

ToR 8. Suggested priorities for future improvements to the stock assessment/projection model.

Priority should be given to reducing the influence of bounds, penalties, priors and weights on likelihoods and ensuring that the current assessment model converges consistently to the global minimum of the objective function within the ranges of feasible values of the various parameters, and demonstrating that such convergence has been achieved. Reweighting of size composition data should then be undertaken, using the Francis (2011) or similar approach and possibly exploring the use of the Dirichlet-multinomial approach. Diagnostic plots showing the extent to which each dataset has been matched by predictions, and the magnitude and patterns of deviations from abundance and size compositions should be produced. Likelihood profiles should be produced to identify any tensions that exist regarding the parameter values favoured by the different datasets. Retrospective and sensitivity analyses should be undertaken.

After ensuring model convergence when fitting and confirming that the model provides predictions that match observed data and appears to represent the known population dynamics of the stock, simulation should be used to generate synthetic data of the forms expected if model assumptions were true, and the assessment model should be applied to the synthetic data to confirm that TCSAM02 estimates of parameters match the values used when generating the synthetic data.

In parallel with the above, Tanner crab data from the paired tow study, i.e. the trawl data from the BSFRF and NMFS vessels from the side-by-side trawling study, should be analysed (outside the assessment model) using mixed effects models with the goal of determining the selectivity of the NMFS bottom trawl. The results from this study should be examined and consideration given as to how best to incorporate the BSFRF side-by-side trawl data into the assessment model.

Recognising that the current side-by-side BSFRF study only fishes for five minutes at the start of the 30 minute tow by the research vessel, consideration should be given as to whether the number of five minute BSFRF tows could be expanded, e.g. with tows at the middle and end of the 30 minute NMFS tow. The additional data would assist in determining the variability of the 5-minute tows.

5. Conclusions and recommendations

Through the use of a control file that allows specification of configuration at run time, the stock assessment model that has been developed for the eastern Bering Sea Tanner crab population, i.e. TCSAM02, provides a very flexible tool to explore alternative hypotheses regarding the population dynamics of this species. The model, which retains consistency with earlier assessment models for this species, has been extended and refined, accommodating different fishing fleets and surveys, and fitting new types of data, e.g. growth data from the eastern Bering Sea. While considerable advances have been made, the model currently fails to converge consistently to the global minimum of the likelihood surface. To overcome this, the best fitting of a number of jitter runs has been accepted as the base case model for use in determining stock status and calculating OFL. Selection of the best fitting of a relatively small number of jitter runs, where key parameters may be trapped at their bounds, does not guarantee that status determinations and estimates of OFLs are reliable, however. Further development of the assessment model is required to remove bounds and penalties, to ensure that priors are not acting as constraints, and to ensure that weighting of likelihoods is appropriate and that the model consistently converges to the global minimum of the likelihood surface. This work should be given priority over other research, as it is required before the model is further extended to incorporate additional survey data or to explore other factors or aspects of the population dynamics.

After overcoming the issues relating to convergence when fitting, and removing (to the extent possible) bounds, penalties, informative priors, and weights on likelihoods, it is recommended that simulation studies be undertaken. These simulations should generate synthetic data under the same assumptions as those employed by TCSAM02. TCSAM02 should then be fitted to the simulated data to assess whether it is able to produce reliable estimates of the parameters that were used to generate the synthetic data. The intent of such analysis is to demonstrate that, when applied to data from the eastern Bering Sea Tanner crab population, the results produced by TCSAM02 are reliable, and to develop an appreciation of the likely precision of the estimates that it produces.

The data from the paired-tow study, i.e. the side-by-side trawling survey undertaken by the BSFRF, are likely to provide valuable information. It is recommended that the Tanner crab data be subjected to a mixed effects analysis, independent of the stock assessment model, with the goal of determining the selectivity of the NMFS bottom trawl. Such analysis is likely to determine how best the data subsequently might be incorporated within TCSAM02.

Research on growth and molting to maturity should continue, recognising that, to the extent possible, the growth data should be representative of the growth of individuals from the entire population, and that factors such as temperature and year might influence growth.

A key area for future research is the need to demonstrate that Mature Male Biomass (MMB) is an appropriate proxy for female reproductive potential. It is possible that the simulation model developed to assess the reliability of TCSAM02 could be extended and employed to explore this issue.

The review process, which has been established by the NMFS, ensures that the data employed and the methods that are used to assess stock status are independently assessed at appropriate intervals. When undertaking an assessment, each Panelist must become familiar with a large number of documents, many of which have been published in earlier years. While it is appropriate that documents relating to the specific subject of the assessment are distributed shortly before the review meeting, there would be value in providing Panelists with access to earlier publications at the time when those Panelists are first appointed such that these may be studied prior to receipt of the final assessment documents.

During the course of review meetings, stock assessment teams are often requested to re-run the assessment models with various options, e.g. creation of a likelihood profile or setting the weight of a likelihood component to zero. It would possibly facilitate communication and documentation of the resulting model output if model results produced during the review used software such as ‘Markdown’, making use of the concepts of ‘reproducible research’, and thereby providing full details of the model configuration used in each analysis.

While placing considerable demands on the individual(s) responsible for the stock assessment, the NMFS review process provides the opportunity for the assessment to be subjected, constructively, to examination by a review panel with ‘fresh eyes’ and experience with models for other fisheries. Such review has the potential to identify issues that have been overlooked, sometimes exposing problems, but also endorsing the approaches that have been employed.

6. Reference list

Cadigan, N. G. and Dowden, J. J., 2010. Statistical inference about the relative efficiency of a new survey protocol, based on paired-tow survey calibration data. *Fishery Bulletin*, **108**: 15-30.

Fowler, G. M. and Showell, M. A., 2009. Calibration of bottom trawl survey vessels: comparative fishing between the Alfred Needler and Teleost on the Scotian Shelf during the summer of 2005 (No. 2824). Fisheries and Oceans Canada.

Francis, R. I. C. C. 2011. Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences*, **68**: 1124–1138.

Kotwicki, S., Lauth, R. R., Williams, K. and Goodman, S. E., 2017. Selectivity ratio: A useful tool for comparing size selectivity of multiple survey gears. *Fisheries Research*, **191**: 76-86.

McAllister, M. K. and J.N. Ianelli, J. N. 1997. Bayesian stock assessment using catch-at-age data and the sampling-importance resampling algorithm. *Can. J. Fish. Aquat. Sci*., **54**: 284–300.

Stevens, B. G., Donaldson, W. E., Hagga, J. A., and Munk, J. E.. 1993. Morphometry and maturity of paired Tanner crabs, *Chionoecetes bairdi*, from shallow and deep water environments. *Canadian Journal of Fisheries and Aquatic Scienc*es **50**: 1504-1516.

Thorson, J. T., Johnson, K. F., Methot, R. D. and Taylor, I. G. 2016. Model-based estimates of effective sample size in stock assessment models using the Dirichlet-multinomial distribution. *Fisheries Research*, http://dx.doi.org/10.1016/j.fishres.2016.06.005.

Appendix 1: Bibliography of materials provided for review

The list below includes material circulated prior to the Review Meeting and also material that was made available during the meeting.

Statement of Work. National Oceanic and Atmospheric Administration (NOAA). National Marine Fisheries Service (NMFS). Center for Independent Experts (CIE) Program. External Independent Peer Review. Bering Sea Tanner Crab Stock Assessment Review. [SoW peer review_TannerCrab.pdf]

Tentative Agenda (2017-06-27). Bering Sea Tanner Crab Stock Assessment Review. [TannerCrab_CIEReviewAgenda.20170627.pdf]

Office of Management and Budget. Memorandum M-05-03. Final Information Quality Bulletin for Peer Review. [OMB_Peer_Review_Bulletin_m05-03.pdf]

Assessment model runs [AssessmentModelRuns folder]

Base model [BaseModel folder]

ModelRun folder: Folder containing input data and model output for base model.

Miscellaneous data files for alternative model runs, together with R code to prepare model output for use with R

Base model plus growth data [BaseModel+GrowthData folder]

ModelRun folder: Folder containing input data and model output for base model.

Miscellaneous data files for alternative model runs, together with R code to prepare model output for use with R

Base model plus high foreign catches [BaseModel+HighForeignCatches folder][**Review request**]

ModelRun folder: Folder containing input data and model output for base model.

Miscellaneous data files for alternative model runs, together with R code to prepare model output for use with R

Base model with no parameter bounds [BaseModelNoLimits folder][**Review request**]

ModelRun folder: Folder containing input data and model output for base model.

Miscellaneous data files for alternative model runs, together with R code to prepare model output for use with R

Final model [FinalModel folder]

ModelRun folder: Folder containing input data and model output for base model.

Miscellaneous data files for alternative model runs, together with R code to prepare model output for use with R

Model comparisons [ModelComparisons folder]

R Markdown code and output for model comparisons and differences for B0, B1, and B2.

R packages [R_Packages folder]

Compressed files containing R code for utilities to process model output.

Notes on model runs [NotesOnModelRuns.docx]

Files containing executable code, i.e. runTCSAM02.pin.bat, runTCSAM02.pin.sh, tcsam02.exe, and tcsam02.osx)

Documents

Assessment documents

Report to May 2017 CPT Meeting [201705ReportToCPT folder]

Stockhausen, W. T. (2017). Tanner Crab Assessment Report for the May 2017 CPT Meeting. Alaska Fisheries Science Center, April 2017.

Appendix A: Corrected Retained Catch Size Frequencies in the Directed Tanner crab Fisheries.

Appendix B: Tanner crab growth (molt increment) data.

Appendix C: Tanner Crab Bycatch in the Groundfish Fisheries.

Appendix D1: Model Comparisons for TCSAM2013 Models B0, B1, B2, B3, B4, B5, B6.

Appendix D2: Model Differences for B1, B2, B3, B4, B5 and B6 vs. B0.

Appendix E: Model Differences between T13B6 and T02A.

Appendix F1a: Model Comparisons for T02A vs AG0.

Appendix F1b: Model Differences for T02A vs AG0.

Appendix F2a: Model Comparisons for AG1 vs AG0.

Appendix F2b: Model Differences for AG1 vs AG0.

Appendix F3a: Model Comparisons for AG1, AG2a, AG2b and AG3.

Appendix F3b: Model Differences for AG1, AG2a, AG2b, and AG3.

Appendix G1: Model Comparisons for TCSAM02 Models AG1, AG1a, AG1b and AG1d.

Appendix G2: Model Differences for TCSAM02 Models AG1a, AG1b, and AG1d vs. AG1.

Appendix H1: Model Comparisons for TCSAM02 Models AG1 and AG1c.

Appendix H2: Model Differences for TCAM02 Models AG1c vs AG1.

Appendix I1: Model Comparisons for TCSAM02 Models AG1 and AG1e.

Appendix J2[*sic*; I2]: Model Differences for TCSAM02 Models AG1 vs AG1e

Appendix J1: Model Comparisons for TCAM02 Models AG3, AG3a, AG3b, and AG4.

Appendix J2: Model Differences for AG3a, AG3b and AG4 vs. AG3.

Appendix K1: Model Comparisons for TCSAM02 Models B1, AG4, and AG1c.

Appendix K2: Model Differences for TCSAM02 Models B1, AG4, and AG1c.

NPFSC BSAI Crab SAFE. (2016). 2016 Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries in the Bering Sea and Aleutian Islands [2016 Introduction Chapter Crab SAFE.pdf]

North Pacific Fishery Management Council's Crab Plan Team. (2016). Crab Plan Team Report. October 2016. [CrabPlanTeamReport_2016-10.pdf]

Punt, A., and Kinzey, D. (prep. by) Report of the Alaska Crab Stock Assessment Workshop, 13-14 May 2009. [SAFE_Guidelines.pdf]

NPFC Scientific and Statistical Committee. (2017). Report to the North Pacific Fishery Management Council, June 5th – 7th, 2017. [SSC Report June 2017_Final.pdf]

Stockhausen, W. T. (2016).2016 Stock Assessment and Fishery Evaluation Report for the Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Regions. September 2016 Plan Team draft. [TannerCrab201609_SAFEReport.pdf]

Stockhausen, W. T. (2017). 2017 Crab Modeling Workshop Report for Tanner Crab. [TannerCrab201701_ModelingWorkshop.pdf]

Anon. (2017). Report of Crab Plan Team Meeting, at Juneau, AK, May 2-5, 2017. [Crab Plan Team meeting May 2017 - Google Docs.pdf]

Assessment model description

Anon. (no date) TCSAM02: The Tanner Crab Stock Assessment Model, version 2. [TCSAM02Description.201707.CIEReview.pdf]

Biology

Donaldson, W. E., Cooney, R. T., and Hilsinger, J. R. (1981). Growth, age and size at maturity of Tanner crab, *Chionoecetes bairdi* M. J. Rathbun, in the northern Gulf of Alaska (Decapoda, Brachyura). *Crustaceana*, 40: 286–302. [DonaldsonEtAl1981_Crustaceana.pdf]

Jadamec, L. S., Donaldson, W. E., and Cullenberg, P. (1999). *Biological Field Techniques for Chionoecetes Crabs*. University of Alaska Sea Grant College Program; Fairbanks, Alaska. [AK-SG-99-02PDF-Biological Field Techniqu.pdf]

Livingston, P. A. (1989). Interannual trends in Pacific Cod, *Gadus macrocephalus*, predation on three commercially important crab species in the eastern Bering Sea. *Fishery Bulletin*, **87**: 807–827. [Livingston1989_FishBull.PredationMortality.pdf]

Paul, A. J., and Paul, J. M. (1995). Molting of functionally mature male *Chionoecetes bairdi rathbun* (decapoda: majidae) and changes in carapace and chela measurements. Journal of Crustacean Biology, **15**: 686–692. [Paul_Paul1995b_JCrustBiol.Growth.pdf]

Siddeek, M. S. M., Zheng, J., Morado, J. F., Kruse, G. H., and Bechtol, W. B. (2010). Effect of bitter crab disease on rebuilding in Alaska Tanner crab stocks. *ICES Journal of Marine Science*, **67**: 2027–2032. [Sideek2010_ICESJMS.BitterCrabDiseaseMortalityForTannerCrab.pdf]

Somerton, D. A. (1981). Regional variation in the size of maturity of two species of tanner crab (*Chionoecetes bairdi* and *C. opilio*) in the eastern Bering Sea, and its use in defining management subareas. *Can. J. Fish. Aquat. Sci*., **38**: 163-174. [Somerton1981_CJFAS.Maturity.pdf]

Stockhausen, W. T. (2013). Appendix to the 2013 BSAI Tanner Crab SAFE Report: Recruitment Analysis for Stock Status Determination and Harvest Recommendations. [TannerCrab2013Appendix_RecruitmentAnalysis.docx; TannerCrab2013Appendix_RecruitmentAnalysis.pdf]

Stockhausen, W. T. (2017). NMFS Survey Data: Chela Height Data and Male Maturity Ogives. [getChelaHeights.docx]

Stoner, A. W., Rose, C. S., Munk, J. E., Hammond, C., and Davis, M. W. (2008). An assessment of discard mortality for two Alaskan crab species, Tanner crab (*Chionoecetes bairdi*) and snow crab (*C. opilio*) based on reflex impairment. *Fish. Bull*., **106**: 337–347. [StonerEtAl_FishBull2008_RAMP.HandlingMortality.pdf]

Tamone, S. L., Taggart, S. J., Andrews, A. G., Mondragon, J., and Nielsen, J. K. (2007). The relationship between circulating ecdysteroids and chela allometry in male

Tanner crabs: evidence for a terminal molt in the genus *Chionoecetes*. *Journal of Crustacean Biology*, **27**: 635–642. [TamoneEtAl2007_JCrustBiol.pdf]

Urban, D., and Hart, D. 1999 Biology of the tanner crab *Chionoecetes bairdi* in Alaska: a report to the Alaska Board of Fisheries. Kodiak, Alaska: Alaska Dept. of Fish and Game, Division of Commercial Fisheries. Regional Information Report No. 4K99-22. [Urban_Hart_RIR.4K.1999.22.pdf]

Zheng, J. (2008). Temporal changes in size at maturity and their implications for fisheries management for eastern Bering Sea Tanner crab. *J. Northw. Alt. Fish. Sci*., **41**: 137–149. [Zheng2008_JNorthwAtlFishSci.Maturity.pdf]

Zheng, J., and Kruse, G. H. (1998). Stock–recruitment relationships for Bristol Bay Tanner crab. *Alaska Fish. Res. Bull.*, **5**: 116–130. [ZhengAndKruse1998_AFRB.StockRecruitRelationships.pdf]

Fisheries

Leon, J. M., Shaishnikoff, J., Nichols, E., and Westphal, M. (2017). *Annual Management Report for Shellfish Fisheries in the Bering Sea–Aleutian Islands Management Area, 2015/16*. Alaska Department of Fish and Game. Fishery Management Report No. 17-10. [ADFG.FisheryManagementReport17-10.pdf]

Alaska Department of Fish and Game. (no date). 2015–2017 King and Tanner Crab Commercial Fishing Regulations. Alaska Department of Fish and Game, Anchorage, Alaska. [ADFG.Regulations_KingTannerCrab-2015-2017.pdf]

Stockhausen, W. T. (2017). *A brief history of the Tanner crab fishery*. Unpublished. [TannerCrabFisheryHistory.pdf]

Management

Staff of the Alaska Fisheries Science Center, Seattle and Kodiak. (2008). *Final Environmental Assessment for Amendment 24 to the Fishery Management Plan for Bering Sea/Aleutian Islands King and Tanner crabs to revise overfishing definitions.* [NPFMC.FinalEAforAmendment24.pdf]

Anon. (2011). *Fishery Management Plan for Bering Sea/Aleutian Islands King and Tanner Crabs.* North Pacific Fishery Management Council, Anchorage, Alaska. [NPFMC.FMP_Crab2011.pdf]

Surveys

NMFS

Daly, B. J., Armistead, C. E., and Foy, R. J. (2016). The 2016 Eastern Bering Sea Continental Shelf Bottom Trawl Survey: Results for Commercial Crab Species. NOAA Technical Memorandum NMFS-AFSC-327. [NOAA-TM-AFSC-327_2016_EBS_ShelfSurvey.pdf]

Somerton, D. A., and Otto, R. S. (1999). Net efficiency of a survey trawl for snow crab, *Chionoecetes opilio*, and Tanner crab, *C. bairdi*. Fish. Bull. **97**:617–625. [Somerton_Otto_FishBull1999.Selectivity.pdf]

Postscript files containing miscellaneous output:

- BottomTempMaps.pdf
- TrawlSurveyData.CPUE.ScaledFemales.SHELL_CONDITION=ALL.MATURITY=IMMATURE.pdf
- TrawlSurveyData.CPUE.ScaledFemales.SHELL_CONDITION=ALL.MATURITY=MATURE.pdf

- TrawlSurveyData.CPUE.scaledImmatureMales.SHELL_CONDITION=NEW_SHELL.SIZE=0.pdf
- TrawlSurveyData.CPUE.scaledLegalMales.SHELL_CONDITION=ALL.SIZE=LEGAL.pdf
- TrawlSurveyData.CPUE.scaledMatureMales.SHELL_CONDITION=ALL.SIZE=MATURE.pdf

BSFRF

Stockhausen, W. T. (2017). BSFRF Side-by-Side Survey Results. [BSFRFSurveys.pdf]

Appendix 2: Copy of the CIE Statement of Work

Statement of Work

National Oceanic and Atmospheric Administration (NOAA)
National Marine Fisheries Service (NMFS)
Center for Independent Experts (CIE) Program
External Independent Peer Review

Bering Sea Tanner Crab Stock Assessment Review

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards. (http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf).
Further information on the CIE program may be obtained from www.ciereviews.org.

Scope

The Alaska Fisheries Science Center (AFSC) Resource Ecology and Fishery Management (REFM) Division requests an independent review of the stock assessment/projection model used to conduct the Bering Sea Tanner crab stock assessment. The model is a size-based integrated assessment model and has been under continuous development since being approved for use by the North Pacific Fisheries Management Council (NPFMC) in 2012. It is anticipated that the North Pacific Fisheries Management Council's Crab Plan Team (CPT) and Science and Statistical Committee (SSC) will approve a change in the TCSAM (Tanner Crab Stock Assessment) code used for the assessment from "TCSAM2013", the code used for the 2013-2016 assessments, to "TCSAM02", a new modeling framework that provides a much more flexible environment than TCSAM2013 for defining alternative models based on a set of model configuration files, as well as fitting new data types not incorporated in TCSAM2013: molt increment (growth) and male chela height (maturity) data. TCSAM02 also calculates the OFL and associated quantities directly within a model run, and thus retains full model uncertainty when using MCMC, whereas using TCSAM2013 the OFL is calculated in a separate projection model and incorporates uncertainty only in recruitment and end-year mature biomass. This review will encompass the TCSAM02 stock assessment/projection model structure and assumptions on which it is based, as well as the life history, fishery, and survey data incorporated in the model. It will also address alternatives for incorporating several industry-funded surveys into the assessment. The Terms of Reference (TORs) for the requested peer review are described in more detail in Annex 2.

Requirements

NMFS requires three (3) CIE reviewers with the necessary qualifications to complete an impartial and independent peer review in accordance with the tasks and TORs (Annex 2) described in the Statement of Work (SOW) herein. The CIE reviewers shall have expertise in conducting stock assessments for fisheries management and be thoroughly familiar with various subject areas involved in stock assessment, including population dynamics, size-structured models, harvest strategies, survey methodology, and the AD Model

Builder programming language to complete the tasks of the scientific peer-review described herein. Familiarity with invertebrate stock assessment, knowledge of crab life history and biology, and harvest strategy development is desirable.

Tasks for reviewers

- Review the following background materials and reports prior to the review meeting:
  1. *Stockhausen, W. 2017. May 2017 Tanner Crab Stock Assessment Activities Report. In prep. [For review:*
  2. *Stram, D. et al. 2016. Introduction Chapter. In: 2016 Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries in the Bering Sea and Aleutian Islands. North Pacific Fisheries Management Council, Anchorage, AK. http://npfmc.legistar.com/gateway.aspx?M=F&ID=2f46b828-51ca-4a45-95bb-cddae2ed8f1d.pdf. [Review the "Stock Status Definitions" and "Status Determination Criteria" for background on the NPFMC's crab stock status criteria and OFL determination]*
  3. *Stockhausen, W. 2016. 2016 Stock Assessment and Fishery Evaluation Report for the Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Regions. In: 2016 Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries in the Bering Sea and Aleutian Islands. North Pacific Fisheries Management Council, Anchorage, AK. http://npfmc.legistar.com/gateway.aspx?M=F&ID=0e48278f-160e-426b-972e-f4736e7c8726.pdf. [The last stock assessment, based on the TCSAM2013 model code.]*
  4. *DALY, B. J., C. E. ARMISTEAD, and R. J. FOY. 2016. The 2016 eastern Bering Sea continental shelf bottom trawl survey: Results for commercial crab species, 167 p. NTIS No. PB2016-104795. [Report on the 2016 NMFS annual eastern Bering Sea shelf summer crab/groundfish trawl survey.]*
  5. *A document (TBD) describing the Gmacs assessment framework.*
  6. *A document (TBD) describing the BSFRF surveys*
- Attend and participate in the panel review meeting:
  - The meeting will consist of presentations by NOAA and other scientists, stock assessment authors and others to facilitate the review, to provide any additional information required by the reviewers, and to answer any questions from reviewers.
- After the review meeting, reviewers shall conduct an independent peer review in accordance with the requirements specified in this SOW, OMB guidelines, and TORs, in adherence with the required formatting and content guidelines; reviewers are not required to reach a consensus.
- Each reviewer may assist the Chair of the meeting with contributions to the summary report, if required by the TORs.
- Deliver their reports to the Government according to the specified milestone dates.

Foreign National Security Clearance

When reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for reviewers who are non-US citizens. For this reason, the reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/ and http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.html. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance

Each CIE reviewer shall participate in, and conduct an independent peer review during, the panel review meeting at the Alaska Fisheries Science Center (AFSC) in Seattle, Washington. Pre- and post-review performance shall be conducted at the contractor's facilities.

Period of Performance
The period of performance shall be from the time of award through September 15, 2017. Each reviewer's duties shall not exceed 14 days to complete all required tasks.

Schedule of Milestones and Deliverables:
The contractor shall complete the tasks and deliverables in accordance with the following schedule.

Within two weeks of award	Contractor selects and confirms reviewers
No later than 17 July 2017	Contractor provides the pre-review documents to the reviewers
31 July – 3 August 2017	Panel review meeting
17 August 2017	Contractor receives draft reports
7 September 2017	Contractor submits final reports to the Government

Applicable Performance Standards
The acceptance of the contract deliverables shall be based on three performance standards:
(1) The reports shall be completed in accordance with the required formatting and content as described in Annex 1; (2) The reports shall address each TOR as specified in Annex 2; (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (http://www.gsa.gov/portal/content/104790). International travel is authorized for this contract. Travel is not to exceed $14,000.

Restricted or Limited Use of Data
The contractors may be required to sign and adhere to a non-disclosure agreement.

NMFS Project Contact
William Stockhausen
william.stockhausen@noaa.gov
National Marine Fisheries Service
7600 Sand Point Way, NE, Bldg. 4,
Seattle, WA 98115-6349
Phone: (206) 526-4241

Annex 1: Peer Review Report Requirements

1. The report must be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether or not the science reviewed is the best scientific information available.

2. The report must contain a background section, description of the individual reviewers' roles in the review activities, summary of findings for each TOR in which the weaknesses and strengths are described, and conclusions and recommendations in accordance with the TORs.

   a. Reviewers must describe in their own words the review activities completed during the panel review meeting, including a brief summary of findings, of the science, conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each TOR even if these were consistent with those of other panelists, but especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the summary report that they believe might require further clarification.

   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

   e. The report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The report shall represent the peer review of each TOR, and shall not simply repeat the contents of the summary report.

3. The report shall include the following appendices:

   1. Appendix 1: Bibliography of materials provided for review
   2. Appendix 2: A copy of this Statement of Work
   3. Appendix 3: Panel membership or other pertinent information from the panel review meeting.

Annex 2: Terms of Reference for the Peer Review

Bering Sea Tanner Crab Stock Assessment Review

The report generated by the consultant should include:

1. Statements assessing the strengths and weaknesses of the current Tanner crab stock *assessment* model with regard to population dynamics, fishery and survey components, likelihood components, and model evaluation.

2. Statements assessing the strengths and weaknesses of the current Tanner crab stock *projection* model, with regard to methodology.

3. A review of the fishery dependent and independent data inputs to the stock assessment with regard to quality of information and appropriateness to the assessment.

4. Recommendations for alternative approaches to evaluate model convergence and compare multiple models.

5. Recommendations for integrating BSFRF surveys into the assessment.

6. Recommendations for alternative assessment/projection model configurations.

7. Recommendations for research that would reduce the uncertainty associated with key parameters assumed or estimated in the assessment.

8. Suggested priorities for future improvements to the stock assessment/projection model.

Annex 3: Tentative Agenda

Bering Sea Tanner Crab Stock Assessment Review

NOAA Alaska Fisheries Science Center
7600 Sand Point Way NE
Seattle, WA 98115

August 2017

Monday, July 31
09:00 Welcome and Introductions
09:15 Role of chair and reviewers, terms of reference
09:30 Overview (fishery, catch levels, bycatch, surveys)
10:30 Biology (growth, natural mortality, maturity curves, mating, molting frequency)
12:00 Lunch
13:00 Survey methodology
14:30 Fishery history and current operation
15:30 Harvest control rules and overfishing definition
17:00 Evening break

Tuesday, Aug. 01
09:00 Stock assessment and projection model
12:00 Lunch
13:00 Stock assessment and projection model (continued)
17:00 Evening break

Wednesday, Aug. 02
9:00 Current research studies
growth, fecundity and egg production
BSFRF side-by-side surveys and other research
12:00 Lunch
1300 Strategies for integrating BSFRF surveys into assessment
14:00 Gmacs
17:00 Evening break

Thursday, Aug. 03
9:00 Reviewer discussions with assessment author.
Review of requested model runs if required.

Appendix 3: Panel membership and participant list

Participants
Stock Assessment Review Panel for
Tanner Crab Stock Assessment Model

NOAA Alaska Fisheries Science Center
7600 Sand Point Way NE
Seattle, WA 98115

July 31 – August 3, 2017

Martin Dorn	Stock assessment chair
Anders Neilsen	CIE Reviewer
Cathy Dichmont	CIE Reviewer
Norman Hall	CIE Reviewer
Buck Stockhausen	Tanner Crab, Pribolof Island King crab and Groundfish stock assessments
Ben Daly	ADFG, Harvest strategy and Tanner Crabs TAC
Scott Goodman	Fisheries Consultant, BSFoundation
Gary Stauffer	Adviser for BSFoundation, Past director of AS modelling
Jack Turnoch	Developer of snow crab model and first Tanner crab assessment.

Webinar participants

Bob Foy	ADFG, Kodiak
Miranda Westphal	ADFG, Dutch Harbor